DOCUMENT RESUME

ED 353 271	TM 018 974

AUTHOR	Wang, Jing-Jen
TITLE	An Analytical Approach to Generating Norms for Skewed Normative Distributions.
PUB DATE	Apr 92
NOTE	14p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Computer Assisted Testing; Computer Oriented Programs; Inferences; Personality Measures; *Scores; *Statistical Distributions; *Test Norms
IDENTIFIERS	*Behavior Assessment System for Children; Minnesota Multiphasic Personality Inventory; *Skew Curves; Smoothing Methods; T Values

ABSTRACT

Approaches for norming skewed test score distributions are reviewed, and an analytical approach is proposed for norming the Behavior Assessment System for Children (BASC). Methods used in constructing norms for several well-known personality instruments that are reviewed include: (1) linear T-scores transformation in the Minnesota Multiphasic Personality Inventory (MMPI); (2) normalized T-scores transformation in the MMPI; (3) the normalized T-scores transformation of T. M. Achenbach; and (4) uniform T-scores transformation in the MMPI-2. In the development of the BASC, the project team decided that both T-score and percentile norms should be developed for clinical interpretation. The curve-fitting procedure of N. L. Johnson and linear T-scores transformation were applied to each of the BASC scales. Results indicate that this approach can be applied effectively to a variety of skewed data sets. The method can be used as an efficient curve-smoothing procedure, and the percentile norms can be generated easily by the computer program. The combined use of percentiles and linear T-scores can minimize problems in clinical inferences. In addition, the proposed method, being computerized and analytical, can save considerable time and reduce error. Two tables and two graphs illustrate the study. There is a 12-item list of references. (SLD)

An Analytical Approach to Generating Norms for Skewed Normative Distributions

Jing-Jen Wang

American Guidance Service
Circle Pines, MN 55014

Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, California, April, 1992.

EDRS

An Analytical Approach to Generating Norms for Skewed Normative Distributions

Introduction

It is well recognized that the normal curve has played a prominent role in the development of the theory of mathematical statistics. Frequently, the latent trait measured by a test instrument is assumed to have an underlying normal distribution in the population from which the normative sample was drawn, and any departures from normality of the observed raw-score distribution are assumed to represent error due to sampling or test-construction problems. Raw-score distributions obtained from normative samples are then transformed into normalized scores such as T-scores (a standardized score with a mean of 50 and a standard deviation of 10). Normalization usually is accomplished by forcing the observed raw-score distribution to be as close as possible to a normal distribution. Through the "area" transformation, irregularities and departures from normality in the raw-score distribution are smoothed out, compressed, or stretched.

One advantage of transforming to normalized scores is the transformed distribution has a well-known form that is easily interpretable. Another advantage is that the normalized scores are "percentile-comparable" across different tests or scales (if normalized and converted to the same mean and standard deviation). It makes sense to normalize scores, however, only if the measured trait has an underlying normal distribution. If the raw-score distribution is highly skewed, then normalization can exaggerate or compress small raw score differences between extreme scores.

Although it is true that most of the latent attributes (especially ability and achievement attributes) can be reasonably assumed to have underlying normal distributions, other psychological traits are known to be skewed in their distributions. For example, the scales in the Minnesota Multiphasic Personality Inventory (MMPI) and the Behavior Assessment System for Children (BASC; Reynolds and Kamphaus, in press) are found to be highly skewed in their score distributions. Table 1 and Figure 1 depict the degree of skewness for scales in the BASC Parent Rating Scales and the MMPI, respectively.

When situations arise in which the assumption of normality is unwarranted, what is the best alternative approach to constructing norms? The objective of this paper is to review several approaches that are available and to propose an analytical approach for norming skewed test-score distributions.

Approaches to Norming Skewed Distributions

Although the statistical characteristics of skewed frequency functions had been investigated since English biometrician Karl Pearson, who developed a system of generalized probability curves and published his "Tables of the Incomplete Gamma Function" in 1922, the application of this family of probability curves has been very limited. Gardner (1950) developed a method by fitting the Pearson Type III Curves to derive an interval scale. Unfortunately, Gardner's scaling approach has not been accepted by practitioners in test development field. In the following session, the methods used in constructing norms for several well-known personality instruments will be reviewed.

Linear T-scores Transformation in the MMPI

In the MMPI scales, the raw scores were linearly transformed to standardized T-scores with a mean of 50 and standard deviation of 10 in the normative groups:

$$\text{T-score} = 10\,[(X - \text{Mean})/SD] + 50, \qquad (1)$$

where X is the raw score, and Mean and SD are the mean and standard deviation of the raw scores in the normative sample.

Linear T-scores are standardized scores but not normalized scores. Since the transformation is linear, the shape of the raw-score distribution is preserved, but the mean (50) and standard deviation (10) are the same across scales regardless of the raw-score differences across scales.

One important advantage of linear T-scores is that the transformation formula can be applied to the whole range of raw scores in the normative distribution, including the highly skewed ones. Unfortunately, linear T-scores derived for scales with differing distribution shapes are not percentile-comparable; this compromises the interpretation of profiles of linear T-scores.

Normalized T-scores Transformation in the MMPI

The most commonly used method for deriving percentile-comparable standardized scores is to force every scale distribution into a normal distribution. The normalization process involves several steps: (1) transform the raw scores to percentiles; (2) find the Z-score in the normal distribution corresponding to each percentile; and (3) transform these Z-scores to T-scores. the normalization process also can be achieved by applying a nonlinear transformation such as square-root or log transformation to the raw-scores, and then applying Equation (1) to the transformed raw-scores to derive T-scores.

Colligan and associates (Colligan, Osborne, Swenson, & Offord, 1984) have criticized the use of linear T-scores for the MMPI, and have proposed replacing linear T-scores with normalized T-scores.

Although normalized transformation brings about percentile comparability, it also can pose some problems. As noted earlier, it makes sense to normalize scores only if the measured trait has an underlying normal distribution. Otherwise, small raw score differences between extreme scores may be exaggerated or compressed by the normalization. Another disadvantage of normalized transformation is that clinically meaningful information on the extreme scores could be lost by the normalization, with consequent loss of validity (Hsu, 1984).

Achenbach's Normalized T-scores Transformation

Achenbach (1991) combined the normalized T-scores transformation and some arbitrary conversion rules to scale the Child Behavior Checklist. For a positively skewed raw-scores distribution, a T score of 50 is assigned to raw scores less than or equal to the 50th percentile. For raw scores between the 50th and the 98th percentile, normalized T-score transformation is applied. For raw scores above the 98th percentile (T=70), linear interpolation is applied to assign T scores from 71 to 100 in as many increments as there were raw scores on a scale.

Achenbach's T-score scale has the same characteristics as normalized T-scores for T scores between 50 and 70. For T-scores above 70, it is an arbitrary scale, which is different from linear T-scores or normalized T-scores.

Uniform T-scores Transformation in the MMPI-2

For the MMPI-2, a uniform T-score transformation was adopted. According to Tellegen and Ben-Porath (in press), "the uniform T-scores are percentile-comparable, yet, unlike normalized T-scores, depart minimally from the familiar linear T-scores."

The development of uniform T-scores involves two major steps. First, a composite or average distribution of the raw scores (or linear T-scores) of the normative samples on the eight clinical scales was derived. This "prototype" distribution shows the expected positive skew (skewness is .7). Second, for each of the eight scales, a polynomial regression formula was derived that transforms the raw scores into scores that approximate the composite T-score values:

$$\text{Uniform T-scores} = B_0 + B_1X + B_2D^2 + B_3D^3, \qquad (2)$$

where X is the raw score to be transformed into the predicted uniform T-score, B_0 is the intercept, and B_1, B_2, and B_3 are

regression weights, and where $D = (C - X)$ if $X < C$, and otherwise $D = 0$, and C equals the value of X corresponding to a composite linear T-score of 60 (85th percentile on the composite distribution).

Uniform T-score transformation is a compromise between linear and normalized T-score transformation. It maintains the positive skewness of the transformed scales, and it minimizes the departure from linear T-scores. Equation (2) indicates that when X approaches C, the transformation becomes linear.

Uniform T-score transformation is an improvement over normalized T-score transformation. However, there is no theoretically driven evidence to show that the slightly skewed prototype distribution is the underlying distribution shape for all the scales. Forcing scales with distinguished degree of skewness to a prototype distribution still can pose some problems.

An Analytical Approach to Norming the BASC

During the development of the BASC, several methods of constructing norms, including the methods mentioned above, were carefully considered. The project team decided that both T-score and percentile norms should be developed for clinical interpretation.

As indicated in Table 1, the degree of skewness across scales in the BASC varies greatly. Since there is no theoretically driven assumption about the shape of the underlying distribution, we assume that the shape of the observed raw-score distribution reflects the underlying distribution in the population. To preserve the shape of the raw-score distribution, both linear T-scores transformation and a Johnson curves fitting procedure were used in developing norms for the BASC scales. T-scores were derived by linear transformation, and percentiles were developed by Johnson Curves translation.

Linear T-Scores Transformation in the BASC

As defined in Equation (1), a linear T-score is an expression of the distance of a raw score from the normative sample's mean raw score, in standard deviation units. A linear T-score transformation assumes that the distances between scale points reflect true differences in degree of problem in a scale, regardless of the shape of the distribution. Because of lack of percentile-comparability in the linear T-score transformation, the BASC norm tables present both T-scores and percentiles for each scale. The intention of providing both T-scores and percentiles in the BASC norm tables is to help clinicians more accurately use the profile to make clinical inferences. As noted

by Tellegen and Ben-Porath (1991), lack of percentile-comparability in linear T-scores "does not compromise strictly actuarial interpretations of profile types". The problems arise only "when clinicians' profile interpretations involve inferences regarding differential psychological characteristics that they associate with elevation differences among the clinical scales". Providing both T-scores and percentiles can avoid the problems associated with clinical inferences in profile interpretations.

<u>Deriving Percentiles in the BASC: Fitting Johnson Curves by Moments</u>

The BASC percentiles have been calculated using an algorithm developed by Hill (1976), Hill, Hill, and Holder (1976), and Roid (1989) that is based on systems of frequency curves described by Johnson (1949).

Johnson (1949) described a system of frequency curves consisting of:

A) the lognormal system: $z = \gamma + \delta \ln (x - \xi), \xi < x,$ (3)
B) the unbounded system: $z = \gamma + \delta \sinh^{-1} ((x - \xi)/\lambda),$ (4)
C) the bounded system: $z = \gamma + \delta \ln ((x - \xi)/(\xi + \lambda - x)),$
$\xi < x < \xi + \lambda,$ (5)

where z is a standardized normal deviate, x is the raw score, and $\gamma, \delta, \xi,$ and λ are parameters in Johnson curves.

Given the mean, standard deviation, skewness, and kurtosis of a scale's observed raw-score frequency distribution, one of the three Johnson curves listed above is selected by the computer program and the parameters are estimated. Using the selected transformation, raw scores are transformed to standard normal variable z-scores. Z-scores are then converted to percentiles using a normal curve function. It is important to note that the Johnson curves transformation is used only to derive the percentile estimates for the BASC scales. The transformed normal deviates are not used to obtain normalized T-scores. The fitted curves preserve the skewness and kurtosis of the observed raw-score distributions, and yield percentile estimates that are smoothed. The Johnson-curve-fitting procedure can be applied to data with a variety of distribution shapes, including those with extremely skewed distributions.

Table 2 is an example of fitting Johnson curves by moments. The Depression scale for the child level of the BASC Parent Rating Scales is used to demonstrate the procedure. The table shows the four moments of the raw-score distribution, the Johnson-curve parameter estimates, and the results of Johnson-curve transformation. The table also shows the normalized T-scores, linear T-scores, and the results of fitting Pearson Type

III Curves. The Pearson Type III curve was fitted using a set of Salvosa's "Tables of Pearson Type III Function" in which areas for curves with unit standard deviations are tabled for skewness values ranging from 0.0 to 1.1 (Salvosa, 1930). The observed and fitted relative frequency distributions are shown in Figure 2. A close look at Figure 2 indicates that the Johnson curve fits the observed data very well.

Discussions and Conclusions

The Johnson-curve-fitting procedure and linear T-scores transformation were applied to each of the BASC scales. The results indicate that the procedure can be applied effectively to a variety of skewed data sets. The method can be used as an efficient curve smoothing procedure, and the percentile norms can be generated easily by the computer program.

It is not surprising that T-scores derived from fitting Pearson Type III Curves with the same degree of skewness are similar to T-scores derived from linear transformation, as indicated in Table 2. The "Tables of Pearson Type III Function" are tabulated with the standard unit as the ordinate, which is similar to the definition of linear T-scores.

Unlike the normalized or uniform T-scores transformation, the Johnson curves transformation method can generate smoothed percentiles and still preserve the shape and characteristics of the observed raw-scores frequency distribution. The combined use of percentiles and linear T-scores in the interpretation of profile can minimize problems in clinical inferences. In addition, the proposed method, being a computerized and analytical procedure, has the promise of saving considerable time and avoiding possible error created by hand calculations in the process of norms development.

References

Achenbach, T.M. (1991). Manual for the Youth Self-Report and 1991 Profile. Burlington, VT: University of Vermont Department of Psychiatry.

Colligan, R.C., Osborne, D., & Offord, K.P. (1980). Linear transformation and the interpretation of MMPI T scores. Journal of clinical Psychology, 36, 162-165.

Colligan, R.C., Osborne, D., Swenson, W.M., & Offord, K.P. (1984). The aging MMPI: Development of contemporary norms. Mayo Clinic Proceedings, 59, 377-390.

Gardner, E.F. (1950). Comments on selected scaling techniques with a description of a new type of scale. Journal of clinical Psychology, 6, 38-43.

Hill, I.D. (1976). Normal-Johnson and Johnson-Normal transformations. Applied Statistics, 45, 183-190.

Hill, I.D., Hill, R., & Holder, R.L. (1976). Algorithm AS 99: Fitting Johnson curves by moments. Applied Statistics, 45, 171-182.

Hsu, L.M. (1984). MMPI T scores: Linear versus normalized. Journal of Consulting and Clinical Psychology, 52, 821-823.

Johnson, N.L. (1949). Systems of frequency curves generated by methods of translation. Biometrika, 36, 149-176.

Reynolds, C.R., & Kamphaus, R.W. (in press). Manual for Behavior Assessment System for Children (BASC). Circle Pines, MN: American Guidance Service.

Roid, G.H. (1989). Programs to fit skewed distributions and generate percentile norms for skewed or kurtotic distributions: Continuous norming with the first four moments. Technical Report No. 89-02, Salem, OR: Assessment Research, Inc..

Salvosa, L.R. (1930). Tables of Pearson's Type III function. Annals of Mathematical Statistics, 1, 191-198 and 1-125.

Tellegen, A., & Ben-Porath, Y.S. (in press). The new uniform T-scores for the MMPI-2: Rationale, derivation and appraisal. Psychological Assessment.

Table 1

Degree of Skewness for Scales in the *Behavior Assessment System for Children* (BASC) Parent Rating Scales, Child Level

Scale	Male	Female	Total
Hyperactivity	0.784	0.720	0.772
Agression	1.008	0.847	1.030
Conduct Problems	1.234	1.122	1.333
Anxiety	0.627	0.546	0.583
Depression	1.032	0.969	1.025
Somatization	1.050	1.052	1.057
Atypicality	1.620	1.480	1.553
Withdrawal	0.769	0.927	0.899
Attention Problems	0.324	0.669	0.495
Adaptability	−0.265	−0.277	−0.270
Social Skills	−0.019	−0.128	−0.072
Leadership	−0.015	−0.041	−0.027

Table 2

An Example of Fitting Johnson Curves by Moments, Derivations of Linear T-Scores, and Fitting Pearson Type III Curves: Depression Scale from the BASC Parent Rating Scales, Child Level

MEAN:	6.77	**TYPE OF JOHNSON CURVE:**	The Bounded System
SD:	4.23	**GAMMA VALUE:**	4.4397
SKEWNESS:	1.03	**DELTA VALUE:**	2.1503
KURTOSIS:	1.61	**LAMBDA VALUE:**	85.2339
		XI VALUE:	-3.5290

Raw Score	Area Under Curve	Z	Percentile Rank	Normalized T-Score	Linear T-Score	Pearson Type III Curve	
						t	T-Score
0	0.010262	-2.32	1	27	34	-1.59	34
1	0.039743	-1.75	4	32	36	-1.37	36
2	0.097165	-1.30	10	37	39	-1.14	39
3	0.180556	-0.91	18	41	41	-0.90	41
4	0.281186	-0.58	28	44	43	-0.66	43
5	0.388469	-0.28	39	47	46	-0.42	46
6	0.493338	-0.02	49	50	48	-0.18	48
7	0.589584	0.23	59	52	51	0.06	51
8	0.673844	0.45	67	55	53	0.30	53
9	0.744983	0.66	74	57	55	0.54	55
10	0.803360	0.85	80	59	58	0.77	58
11	0.850183	1.04	85	60	60	1.01	60
12	0.887047	1.21	89	62	62	1.24	62
13	0.915626	1.38	92	64	65	1.48	65
14	0.937494	1.53	94	65	67	1.71	67
15	0.954041	1.69	95	67	69	1.94	69
16	0.966440	1.83	97	68	72	2.17	72
17	0.975651	1.97	98	70	74	2.40	74
18	0.982441	2.11	98	71	77	2.63	76
19	0.987410	2.24	99	72	79	2.86	79
20	0.991024	2.37	99	74	81	3.10	81
21	0.993635	2.49	99	75	84	3.33	83
22	0.995511	2.81	99	78	86	3.56	86
23	0.996850	2.73	99	77	88	3.79	88
24	0.997802	2.85	99	78	91	4.03	90
25	0.998475	2.96	99	80	93	4.26	93
26	0.998947	3.08	99	81	95	4.50	95
27	0.999277	3.19	99	82	98	4.74	97
28	0.999507	3.29	99	83	100	4.98	100
29	0.999666	3.40	99	84	103	5.22	102
30	0.999775	3.51	99	85	105	5.46	105
31	0.999849	3.61	99	86	107	5.70	107
32	0.999899	3.72	99	87	110	5.95	110
33	0.999934	3.82	99	88	112	6.21	112
34	0.999956	3.92	99	89	114	6.45	114
35	0.999972	4.03	99	90	117	6.72	117
36	0.999982	4.13	99	91	119	6.98	120

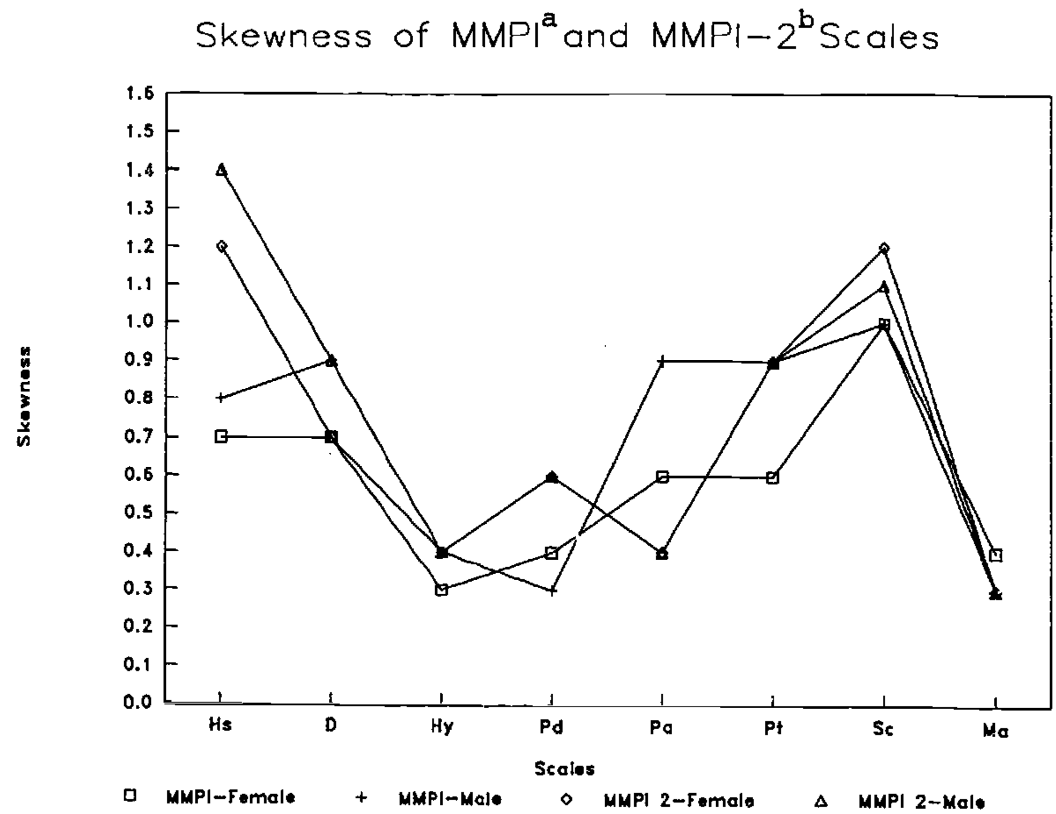

Figure 1: Degree of Skewness of MMPI and MMPI-2 Scales

a The MMPI skewness data obtained from Colligan, Osborne, & Offord, 1980.

b The MMPI-2 skewness data as described by Tellegen and Ben-Porath (in press) in Figure 3.

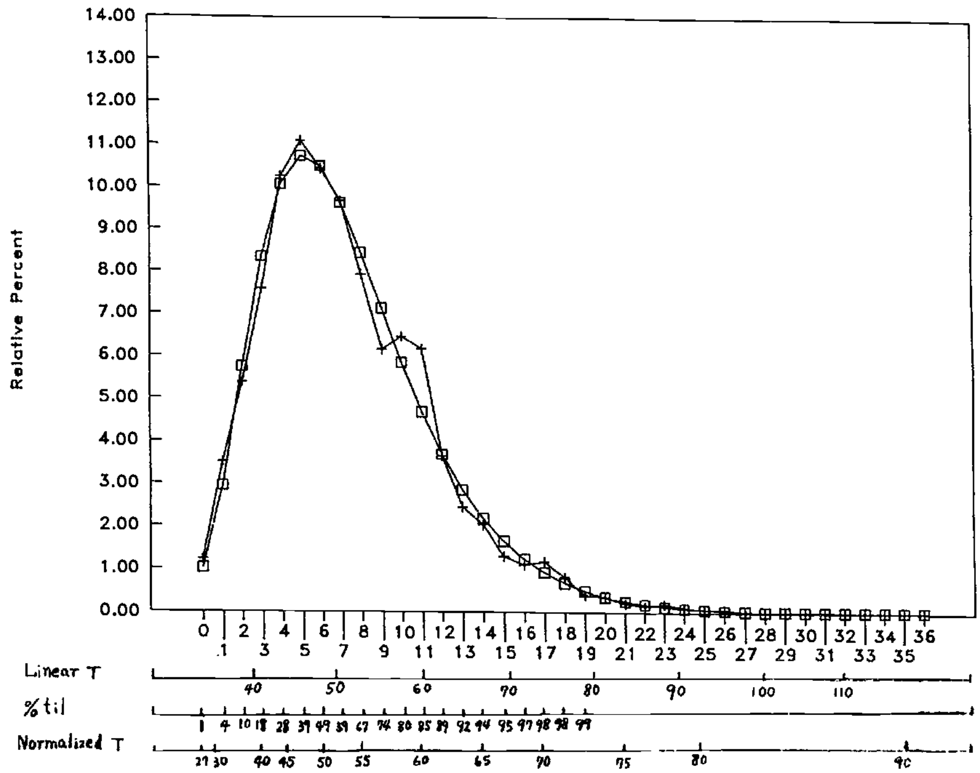

Figure 2: Observed and Smoothed Relative Frequency Distributions for Depression Scale of the BASC Parent Rating Scales, Child Level